



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/067,550	02/05/2002	Ramesh Jain	VIRAGE.023C2	5888

20995 7590 09/18/2006

KNOBBE MARTENS OLSON & BEAR LLP
2040 MAIN STREET
FOURTEENTH FLOOR
IRVINE, CA 92614

EXAMINER

ONUAKU, CHRISTOPHER O

ART UNIT	PAPER NUMBER
2621	

DATE MAILED: 09/18/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary

Application No.

10/067,550

Applicant(s)

JAIN ET AL.

Examiner

Christopher Onuaku

Art Unit

2621

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --
Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☐ Responsive to communication(s) filed on ____.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 1-35 is/are pending in the application.
- 4a) Of the above claim(s) ____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) ____ is/are allowed.
- 6) ☒ Claim(s) 1-35 is/are rejected.
- 7) ☐ Claim(s) ____ is/are objected to.
- 8) ☐ Claim(s) ____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☒ The drawing(s) filed on 05 February 2002 is/are: a) ☒ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
- ☐ Certified copies of the priority documents have been received.
 - ☐ Certified copies of the priority documents have been received in Application No. ____.
 - ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- ☒ Notice of References Cited (PTO-892)
- ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
- ☒ Information Disclosure Statement(s) (PTO/SB/08)
Paper No(s)/Mail Date 2/28/06; 8/22/05; 8/8/05; 6/25/02.
- ☐ Interview Summary (PTO-413)
Paper No(s)/Mail Date. ____.
- ☐ Notice of Informal Patent Application
- ☐ Other: ____.

DETAILED ACTION

Claim Rejections - 35 USC § 102

1. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(e) the invention was described in (1) an application for patent, published under section 122(b), by another filed in the United States before the invention by the applicant for patent or (2) a patent granted on an application for patent by another filed in the United States before the invention by the applicant for patent, except that an international application filed under the treaty defined in section 351(a) shall have the effects for purposes of this subsection of an application filed in the United States only if the international application designated the United States and was published under Article 21(2) of such treaty in the English language.

2. Claims 10,13,15,16,18,20,22,23,27,30,32&33 are rejected under 35 U.S.C. 102(e) as being anticipated by Chow et al (US 5,692,104).

Regarding claim 10, Chow et al disclose method and apparatus for detecting end points of speech activity, comprising:

- a) an audio signal switch receiving an audio signal (speech detection block 230 of speech feature extraction 210 (see Fig.2&3, col.6, line 58 to col.7, line 19);
- b) an audio classification component controlling the audio signal switch according to whether the audio is classified as speech (see Fig.2&3; VQ distortion processing block 303 of speech activity detection block 230; col.7, lines 30-45);
- c) a plurality of audio metadata track extraction components in data communication with the output speech, wherein each audio metadata track extraction

Art Unit: 2621

component provides an audio metadata track associated with speech (see Fig.1&2; speech feature extraction 210; col.5, lines 39-60).

Regarding claim 13, Chow discloses wherein the audio classification component additionally classifies at least silence and music (see col.6, lines 58-65).

Regarding claim 15, Chow discloses wherein the audio signal is received from a real-time source (see microphone disclosed in col.5, lines 16-19).

Regarding claim 16, Chow discloses wherein the audio signal is received from a digital source (see Fig.1; sound sampling device 125; col.5, lines 13-18).

Regarding claim 18, the claimed limitations of claim 18 are accommodated in the discussions of claim 10 above.

Regarding claim 20, the claimed limitations of claim 20 are accommodated in the discussions of claim 13 above.

Regarding claim 22, the claimed limitations of claim 22 are accommodated in the discussions of claim 15 above. Here examiner reads the microphone as a remote real-time audio signal source.

Art Unit: 2621

Regarding claim 23, the claimed limitations of claim 23 are accommodated in the discussions of claim 16 above.

Regarding claim 27, the claimed limitations of claim 27 are accommodated in the discussions of claim 10 above, including the claimed audio class dictionary configured to provide dictionary data indicative of audio classes to the audio classification engine (see Fig.2; recognizer process 220 which performs speech recognition using language model to determine whether the extracted features represent extracted words in a vocabulary recognizable by the speech recognition system; col.6, lines 32-46).

Regarding claim 30, the claimed limitations of claim 30 are accommodated in the discussions of claim 13 above.

Regarding claim 32, the claimed limitations of claim 32 are accommodated in the discussions of claim 15 above.

Regarding claim 33, the claimed limitations of claim 33 are accommodated in the discussions of claim 16 above.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

Art Unit: 2621

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

4. Claims 1-3,5-7,9&24-26 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chang et al (US 5,828,809) in view of Yoshio et al (US 6,034,942).

Regarding claim 1, Chang et al disclose context-based video indexing and video information extraction systems including an information extraction system that combines and integrates both speech understanding and image analysis, comprising the method of:

a) receiving video information having embedded audio information and associated time information (see col.3, lines 20-58 and col.5, lines 13-22);

b) capturing the embedded audio information in the video information (see Fig.2, col.3, line 59 to col.4, line 10), here the audio and video components are separated and separately digitized;

c) extracting a plurality of audio metadata tracks from the audio information, each audio metadata track having selected ones of the time information indicative at least of start and stop times for the audio metadata track, encoding the video information, and accessing the encoded video information with the selected time information of one of the audio metadata tracks (see Fig.2; Audio Signal Analysis; Video Information Analysis and the wordspotting algorithm; col.3, line 64 to col.6, line 44), here the video data and the audio data are digitized, audio analysis module locates candidates in the digitized audio data by performing wordspotting. This information is passed to the video analysis module 66 which analyzes the video data by segmenting and identifying the shots. The

Art Unit: 2621

indexing information from the video analysis module 66 is in the form of pointers to the locations of interesting events. For example, in sports programs, information content in audio is highly correlated with the information content in video. In the game of football, for example, important keywords such as "touchdown" or "fumble" can be detected in the audio stream, and this audio data can be used as a coarse filter to locate candidates for important events. In video analysis, assuming that a touchdown candidate is located at a time t , video analysis is applied to the region of $t-1$ minutes and $t+2$ minutes, the assumption being that a touchdown event should begin and end within that time range. In video processing, the original video sequence is broken down into discrete shots. Key frames from each shot are extracted and shot identification is then applied on them to verify the existence of a touchdown. It is pertinent to point out that the examiner reads, for example, the data representing sounds, pointers (or indexes) to the locations, and the locations, where the specific words (e.g., keywords) are found, as examples of audio metadata tracks.

Chang fails to explicitly disclose time information as time codes. Yoshio et al teach recording method for recording the information (video information, audio information, and the like) onto high density information recording medium, and reproducing method for reproducing the information from the information record medium wherein the video and audio data are divided on the basis of time codes (see col.19, line 8 to col.20, line 67). Time codes are well known as means for identifying the start and end of audio and video data during the recording process, for example, in order to facilitate the reproduction process. It would have been obvious to modify Chang by

Art Unit: 2621

adding time code processing means to Chang, as taught by Yoshio, since time codes are well known as means for identifying the start and end of audio and video data during the recording process, for example, in order to facilitate the reproduction process. With Chang modified with Yoshio, it would have been obvious to use time codes for indicating the start and end times for the audio data (metadata) in Chang, during the recording process, for example, in order to facilitate the reproduction process.

Regarding claim 2, Chang discloses the method wherein the video information is received from an analog source (see col.3, lines 64-66).

Regarding claim 3, Chang discloses the method wherein the analog source is a videotape deck (see col.3, lines 64-66).

Regarding claim 5, Chang discloses the method wherein the video information is received from a digital source (see col.3, lines 64-66).

Regarding claim 6, Chang discloses the method wherein the capturing includes digitizing with an audio digitizing devices (see col.3, line 64 to col.4, line 3).

Regarding claim 7, Chang discloses the method wherein the plurality of audio metadata tracks includes at least one of keywords, speech-to-text transcriptions, speaker identification and audio (see col.4, lines 19-27).

Art Unit: 2621

Regarding claim 9, Chang discloses the method wherein the encoding comprises encoding with an MPEG format (see col.3, line 64 to col.4, line 3).

Regarding claim 24, the claimed limitations of claim 24 are accommodated in the discussions of claim 1 above.

Regarding claim 25, the claimed limitations of claim 25 are accommodated in the discussions of claim 5 above.

Regarding claim 26, the claimed limitations of claim 26 are accommodated in the discussions of claim 7 above.

5. Claim 4 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chang et al in view of Yoshio et al and further in view of Yamashita (US 5,963,702).

Regarding claim 4, Chang and Yoshio fail to explicitly disclose satellite as a source of video information. Yamashita teaches using satellite broadcast tuners for receiving information (see Fig.5, and col.8, lines 34-47). Using satellite receivers provides an additional source of video information. It would have been obvious to further modify Chang by realizing Chang with a satellite receiving means, as taught by Yamashita, since this provides an additional source of video information, thereby increasing the dynamic range of Chang.

Art Unit: 2621

6. Claim 8 is rejected under 35 U.S.C. 103(a) as being unpatentable over Chang et al in view of Yoshio et al and further in view of Reichek et al (US 5,701,153).

Regarding claim 8, Chang and Yoshio fail to explicitly disclose wherein the time codes comprise SMPTE codes. Reichek teaches generation and use of SMPTE time codes (see col.6, lines 10-27). It would have been obvious to add SMPTE time codes generation and use capability to Chang, as taught by Reichek, so that Chang can generate and use SMPTE time codes.

7. Claims 11&28 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chow et al in view of Miyamori et al (US 5,677,994).

Regarding claim 11, Chow further discloses an audio capture component for capturing and digitizing an analog audio source (see sound sampling device 125 of the system 100; col.5, lines 13-20). Chow fails to explicitly disclose an audio signal normalization component for normalizing the digitized audio prior to processing. Miyamori et al teach a high-efficiency encoding and decoding method and apparatus in encoding and decoding multi-channel data comprising a block floating unit processor C14 which normalizes the audio data of the respective frequency bands resulting from resolution into the signal components on the block floating unit by the MDCT unit C13. Normalizing audio signals provides the desirable advantage of, for example, bringing the amplitude of the audio spectral envelope to a predetermined level to bring the amplitude of the video signal to a desired level. It would have been obvious to modify Chow by realizing Chow with audio normalizing means, as taught by Miyamori, since

Art Unit: 2621

this provides the desirable advantage of, for example, bringing the amplitude of the audio spectral envelope to a predetermined level to bring the amplitude of the audio signal to a desired level.

Regarding claim 28, the claimed limitations of claim 28 are accommodated in the discussions of claim 11 above.

8. Claims 12, 19 & 29 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chow et al in view of Reichek et al (US 5,701,153).

Regarding claim 12, Chow fails to explicitly disclose wherein the plurality of audio metadata tracks includes at least one of keywords, speech-to-text transcription, speaker identification and audio class. Reichek teaches using keyword search in the selection of a position in a text (see col.13, lines 7-15). Keyword identification capability provides the desirable advantage of identifying keywords which facilitates the classification of audio as speech. It would have been obvious to further modify Chow by realizing Chow with a keyword identification capability, as taught by Reichek, since this provides the desirable advantage of identifying keywords which facilitates the classification of audio as speech.

Regarding claim 19, Reichek further teaches wherein the audio metadata tracks includes at least speaker identification (see col.3, lines 48-57 and col.9, lines 39-56).

Art Unit: 2621

Regarding claim 29, the claimed limitations of claim 29 are accommodated in the discussions of claim 12 above.

9. Claims 14,21,31&35 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chow et al.

Regarding claim 14, Chow fails to explicitly disclose wherein the audio metadata track extraction components receive data from a customizable dictionary, but this would have been an obvious engineering design consideration depending on the circuit at hand.

Regarding claim 21, the claimed limitations of claim 21 are accommodated in the discussions of claim 14 above.

Regarding claim 31, the claimed limitations of claim 31 are accommodated in the discussions of claim 14 above.

Regarding claim 35, the claimed limitations of claim 35 are accommodated in the discussions of claim 14 above.

10. Claims 17&34 are rejected under 35 U.S.C. 103(a) as being unpatentable over Chow et al in view of Cruz et al (US 5,613,032).

Art Unit: 2621

Regarding claim 17, Chow fails to explicitly disclose wherein audio signal is received from a digital camcorder. Cruz teaches a digital camcorder as audio signal source (see Fig.1,2&3A-3C; col.3, lines 30-58; col.10, line 33 to col.11, line 38). Using a digital camcorder provides an additional source of audio signal. It would have been obvious to add a digital camcorder capability to Chow, as taught by Cruz, which would provide an additional audio source means to Chow.

Regarding claim 34, the claimed limitations of claim 34 are accommodated in the discussions of claim 17 above.

Conclusion

11. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Kovalick et al (US 5,485,553) teach video printing, including printing video images on a printable medium.

Setagawa et al (US 5,822,024) teach an image coding method/apparatus, including a coding/decoding method of a motion picture and an apparatus for the same.

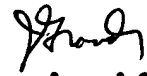
12. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Christopher Onuaku whose telephone number is 571-272-7379. The examiner can normally be reached on M-F.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, James Groody can be reached on 571-272-7950. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Art Unit: 2621

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.


COO
9/13/06


James J. Groody
Supervisory Patent Examiner
Art Unit 262-2621